

# Approximating the Partially Coherent Additive White Gaussian Noise Channel in Polar Coordinates

Martí Sales-Llopis and Seb J. Savory, *Fellow, IEEE*

**Abstract**—We consider the partially coherent additive white Gaussian noise channel (PCAWGN) in optical communications and review the derivation of the exact channel conditional probability model in a closed-form solution in polar coordinates. Additionally, we derive a reduced-complexity approximation by replacing the Rician and Tikhonov distributions describing amplitude and phase components, respectively, by their Gaussian approximation under certain assumptions of high SNR and low phase noise or *jitter*. Our proposal significantly reduces the hardware complexity by removing the modified Bessel functions involved in the exact solution. Furthermore, we compare the proposed approximation with a different metric previously found in the literature and observe that for maximum-likelihood hard symbol decision both models are in perfect agreement with the optimal detector. However, our model not only reduces the required number of multiplications from 12 to 8 and additions from 9 to 3 (per computed symbol), but also reduces the information loss by at least one and up to several orders of magnitude with respect to the previously published metric when used to compute the channel achievable information rate (AIR). In all the simulation cases we use QAM constellations of order 4, 8, 16, and 32 as test input symbol sets.

**Index Terms**—partially coherent AWGN channel, optical communications, phase noise, channel conditional probability

## I. INTRODUCTION

In optical communications it is customary to model different sources of noise, such as the propagated amplified spontaneous emission (ASE) or the shot-noise at the photo-diodes (PD), as a lumped term of additive white Gaussian (complex) noise (AWGN) at the receiver. In addition to that, when the information is not encoded only in the amplitude but also in the phase, coherent optical systems suffer from another kind of angular noise known as phase noise or *jitter*, mainly (but not only) due to the lasers spectral linewidth, where the phase evolution of the laser output signal is commonly modeled as a Wiener process (or random walk), i.e. a process with memory.

One particular way to study such optical channels consists of considering that one or more subsystems at the receiver have performed the required signal conditioning (filtering, equalization, etc.) and the subsequent carrier phase estimation (CPE), either with a phase-locked loop (PLL) or with digital signal processing (DSP) techniques. Thus, what is left is the transmitted signal plus a component of linear noise and a component of angular or phase noise. This is commonly known as the (memoryless) partially coherent AWGN channel (PCAWGN) and it has been widely studied in the past [1]–[4].

The mathematical model of the PCAWGN channel conditional probability can serve more than one purpose. One of utmost importance in modern communications is that it enables the computation of mutual information (MI) as an upper limit to the channel's achievable information rate (AIR) [5]. Among other possible applications, e.g. in [1], [2] the authors used the model on the estimation of the bit-to-error ratio (BER), in [4] it was used to derive a lower bound of 2.4 dB on the shaping gain, and in [6] it was shown that it can improve a CPE algorithm tolerance to phase noise.

In this work we aim to introduce a reduced-complexity approximated model for the PCAWGN channel conditional probability and compare it with other existing solutions in the literature. This letter is structured as follows: Section II provides the mathematical model of the PCAWGN channel conditional probability. Then, in Section III we describe some ways to numerically solve the exact solution and propose a reduced-complexity approximation in polar coordinates. We also include, by way of comparison, another approach previously presented in [1]. In Section IV we present the simulation results of the maximum-likelihood hard-decision (ML-HD) symbol error rate (SER) and information loss for different values of SNR and phase noise. Finally, we end this letter with the conclusions.

## II. PROBLEM STATEMENT

Consider the discrete-time memoryless PCAWGN channel in which at any given sampling time  $t = kT$  we have

$$Y = X \cdot e^{j\varphi} + N \quad (1)$$

where  $X$  and  $Y$  are complex-valued random variables representing the transmitted symbols and the received samples, respectively.  $N$  is complex additive white Gaussian noise (AWGN) with zero mean and variance  $N_0$ , i.e.  $N \in \mathbb{C}$ ,  $\Re(N) \sim \mathcal{N}(0, \frac{N_0}{2})$ , and  $\Im(N) \sim \mathcal{N}(0, \frac{N_0}{2})$ . Each symbol in  $X$  is drawn from the discrete constellation  $\mathcal{X}$  with probability mass function (PMF)  $p(x)$ , and the observed variable  $Y : (\rho, \theta)$  is continuous, with amplitude  $\rho \in [0, \infty)$  (although in practice this range is generally limited to  $[0, \rho_{\max})$ ) and argument  $\theta \in [-\pi, \pi)$ . The term  $\varphi$  represents the phase noise,  $j = \sqrt{-1}$ , and here we define the SNR in dB as  $10 \cdot \log_{10}(P_S/N_0)$  where  $P_S$  is the average transmitted power, i.e.  $P_s = \mathbb{E}[|X|^2]$ , where  $\mathbb{E}$  means statistical expectation.

### A. Special case: the AWGN channel

We find convenient to start considering the channel in the absence of phase noise because it is simpler to first analyze

The authors are with the Electrical Engineering Division, Department of Engineering, University of Cambridge, Cambridge CB3 0FA, U.K. (e-mail: ms2345@cam.ac.uk; sjs1001@cam.ac.uk).

the distributions of amplitude and angular components due to AWGN only, and later introduce the effects of phase noise. Thus, assuming  $\varphi = 0$  in Eq. (1) the channel conditional probability  $p(y|x)$  in polar coordinates is well known as [7]

$$p(y|x) = \frac{\rho}{\pi N_0} \exp\left(-\frac{|x|^2 + \rho^2 - 2|x|\rho \cos \Delta\theta_0}{N_0}\right) \quad (2)$$

where  $\Delta\theta_0$  is the modulo- $\pi$  angular difference between  $x$  and  $y$ , that is

$$\Delta\theta_0 = \theta - \arg(x), \quad \Delta\theta_0 \in [-\pi, \pi). \quad (3)$$

The marginal probability distribution function (PDF) of the observed amplitude  $\rho$  is also well known to follow the Rician distribution

$$p(\rho|x) = \frac{\rho}{N_0/2} \exp\left(-\frac{|x|^2 + \rho^2}{N_0}\right) I_0(\kappa) \quad (4)$$

where  $I_0(\kappa)$  is the modified Bessel function of the first kind and  $\kappa$  is given by

$$\kappa = \frac{|x|\rho}{N_0/2}. \quad (5)$$

From Eq. (2) and Eq. (4) it is possible to obtain the *a posteriori* probability (APP) of the angular difference  $\Delta\theta_0$  as

$$p(\Delta\theta_0|x, \rho) = \frac{\exp[\kappa \cos \Delta\theta_0]}{2\pi I_0(\kappa)}, \quad (6)$$

which is a Tikhonov distribution,  $\mathcal{M}(0, \kappa)$ , (also known as *von Mises*). Eq. (6) is the exact solution for the PDF of the observed AWGN-induced difference  $\Delta\theta_0$  conditioned on the transmitted symbol  $x$  and the observed amplitude  $\rho$  [8].

### B. Study case: the PCAWGN channel

To model the PCAWGN channel conditional probability it is necessary to know the statistical properties of  $\varphi$ . The most commonly used distributions for phase noise in PCAWGN channels are the wrapped Gaussian and the Tikhonov, with zero mean and variance  $\sigma_p^2$ . The former might result impractical due to the infinite series involved in its computation, although a truncated version might be used. On the other hand, the latter is a very popular choice as it is a good approximation for the wrapped Gaussian and can be computed efficiently. For this reason, in what follows we will assume  $\varphi \sim \mathcal{M}(0, \alpha)$ .

Modeling in polar coordinates allows us to treat the amplitude and phase separately: on one hand, the amplitude of the received signal is not affected by the phase noise, hence the distribution  $p(\rho|x)$  is the same as given in Eq. (4) and, on the other hand, the new angular difference  $\Delta\theta = \Delta\theta_0 + \varphi$  has a distribution that results from the convolution of  $p(\Delta\theta_0|x, \rho)$  with  $p(\varphi)$  and is calculated as [9, Eq. 3.5.43]

$$p(\Delta\theta|x, \rho) = \frac{I_0\left(\sqrt{\kappa^2 + \alpha^2 + 2\kappa\alpha \cos(\Delta\theta)}\right)}{2\pi I_0(\kappa) I_0(\alpha)}, \quad (7)$$

with  $\Delta\theta$  defined as in Eq. (3). Multiplying Eq. (4) and Eq. (7) yields the PCAWGN channel conditional probability

$$p(y|x) = \frac{\rho}{\pi N_0} \exp\left(-\frac{|x|^2 + \rho^2}{N_0}\right) \cdot \frac{I_0(\nu)}{I_0(\alpha)} \quad (8)$$

with

$$\nu = \sqrt{\kappa^2 + \alpha^2 + 2\kappa\alpha \cos(\Delta\theta)} \quad (9)$$

which is a similar result to the one shown in [4, Eq. 3] with the only difference that in the latter the authors considered other factors like the portion of power in a pilot tone and the one-sided loop bandwidth-sampling period product. For the remainder of this text we will refer to Eq. (8) as to the *exact* solution of the PCAWGN channel conditional probability.

## III. APPROXIMATED MODELS

### A. Numerical approximations

Numerical approximations are sometimes required to evaluate Eq. (8) due to computational overflows associated with  $I_0(z)$ . This can, however, be overcome by using the asymptotic approximation of  $I_0(z)$  such as the truncated power expansion (see e.g. [9, Eq. 3.5.33]) to give

$$\log p(y|x) \approx \log \frac{\rho}{\pi N_0} - \frac{|x|^2 + \rho^2}{N_0} + \log \hat{I}_0(\nu) - \log \hat{I}_0(\alpha) \quad (10)$$

where

$$\hat{I}_0(z) = \frac{e^z}{\sqrt{2\pi z}} \left\{ 1 + \frac{1}{8z} + \frac{9}{128z^2} + \frac{225}{3072z^3} \right\} \quad (11)$$

### B. Foschini's jitter distance

An approximated model for the PCAWGN, the so-called *jitter distance* by Foschini *et al.*, is defined in [1, Eq. 13] as

$$p(y|x) \approx \frac{1}{2\pi N_0} \exp\left[-\frac{1}{2N_0} d^2(y, x)\right] \quad (12)$$

where the *effective* squared distance  $d^2$  is obtained as

$$d^2(y, x) \triangleq |y - x|^2 + 2\langle y, x \rangle + 2\alpha N_0 - 2\left|\sqrt{\rho^2 |x|^2 + 2\alpha N_0 \langle y, x \rangle + (\alpha N_0)^2}\right| \quad (13)$$

and the angled brackets  $\langle \cdot, \cdot \rangle$  indicate the *inner* product. In what follows we will refer to Eq. (12) and Eq. (13) as Foschini's metric, model, or approximation.

### C. Proposed approximated model in polar coordinates

We derive an alternative metric following a different approach and approximating the channel model of Eq. (8) in polar coordinates. Let us start with the amplitude component: the Rician distribution in Eq. (4) approaches a Gaussian distribution  $\sim \mathcal{N}(|x|, \frac{N_0}{2})$ , for high SNR, since  $\sqrt{\rho/|x|} \approx 1$  and  $I_0(z) \approx \exp(z)/\sqrt{2\pi z}$  giving,

$$p(\rho|x) \approx \frac{K_\rho}{\sqrt{\pi N_0}} \exp\left(-\frac{(\rho - |x|)^2}{N_0}\right) \quad (14)$$

where the constant  $K_\rho$  is to ensure that the total probability between  $\rho \in [0, \infty)$  is unity. Similarly, the phase distribution given by Eq. (6) as the convolution of two Tikhonov can be approximated by another Tikhonov or by a wrapped Gaussian ([9, p. 44]) which we will in turn approximate by truncation to the unwrapped Gaussian

$$p(\Delta\theta|x, \rho) \approx \frac{K_\theta}{\sigma_\Theta \sqrt{2\pi}} \exp\left(-\frac{\Delta\theta^2}{2\sigma_\Theta^2}\right) \quad (15)$$

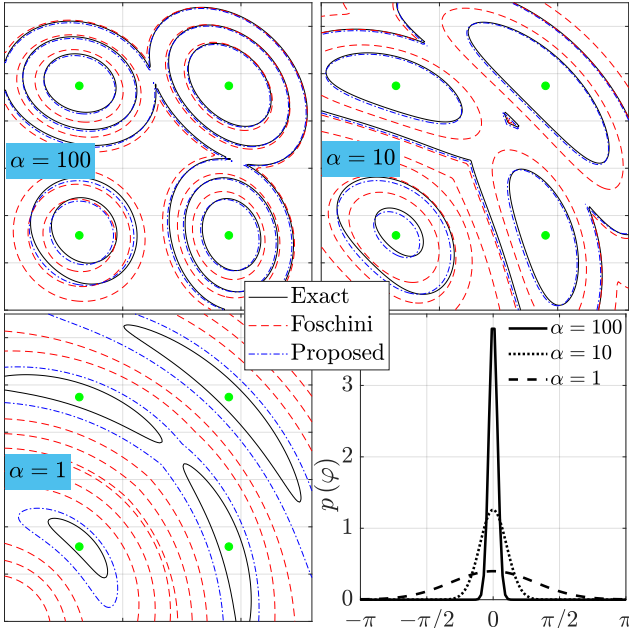


Fig. 1. Contours for equal probability density values (referred to as equiprobable curves) from points in the quadrant I for 16-QAM. The SNR is 15 dB and  $\alpha$  takes values 100, 10, and 1. The curves are obtained with the exact model (solid), Foschini's metric (dashed), and the proposed model (dash-dotted). The PDFs for the phase noise  $\varphi$  are shown for reference.

where  $K_\theta$  ensures the total probability in  $\Delta\theta \in (-\pi, \pi)$  is unity and  $\sigma_\Theta^2$  is the variance of  $\Delta\theta$ , which we obtain as the sum of the variances of the distributions convolved, that is

$$\sigma_\Theta^2 = \kappa^{-1} + \alpha^{-1} = \frac{N_0}{2|x|\rho} + \sigma_p^2 \quad (16)$$

where we have assumed that for  $\sigma_p^2 \ll 1 \rightarrow \sigma_p^2 \approx \frac{1}{\alpha}$ . According to our approximation, half of the power of the AWGN term will produce a Gaussian-distributed angular displacement proportionally shorter as the point is further from the origin. In this regard, the factor  $(|x|\rho)$  is known to yield a better estimation than  $|x|^2$  or  $\rho^2$  [8].

Finally, multiplying Eq. (15) and Eq. (14) we obtain the proposed approximated model

$$p(y|x) \approx \frac{K}{\pi\gamma N_0} \exp\left(-\frac{\Delta\rho^2 + (\gamma\Delta\theta)^2}{N_0}\right) \quad (17)$$

with

$$\Delta\rho = \rho - |x|$$

$$\gamma = \left(\frac{1}{|x|\rho} + \frac{2}{\alpha N_0}\right)^{-\frac{1}{2}} \quad (18)$$

and  $K = K_\rho K_\theta$  which for high SNR we assume  $K \approx 1$ . It is noticeable the way in which Eq. (17) resembles an Euclidean distance. Similarly to what happens with Foschini's metric in Eq. (12), the angles are "compressed" such that points at a longer angular distance appear to be equidistant. These equidistant (or equiprobable) curves are shown in Fig. 1 for three values of probability densities (5, 2.5, and 1). The approximated metrics exhibit the same shape as the exact

TABLE I  
COMPLEXITY COMPARISON OF THE PCAWGN CHANNEL MODELS IN TERMS OF NUMBER OF OPERATIONS REQUIRED PER COMPUTED SYMBOL

Model	+	×	√	exp	>>*	$I_p(z)$
Exact	2	7	1	1	1	2
Foschini	9	12	1	1	2	0
Proposed	3	8	1	1	0	0

\* bit shift-register

model, but only the proposed one closely matches the probability density values for most values of  $\alpha$ . This explains why while both approximations are good as ML-HD detectors (see Fig. 2), our model yields a better calculation of channel MI (see Fig. 3). The main difference between the proposed model and Foschini's metric is the use of polar versus rectangular coordinates, respectively. The latter is derived by means of vector algebra and thus relies on inner products, which require a few more multiplications per symbol (12) compared to the proposed polar model (8). The required number of operations for the approximated and exact solutions is detailed in Table I.

#### IV. SIMULATION RESULTS

In a first set of simulations, we test Foschini's metric and the proposed approximation as ML-HD detectors and compare their performance in terms SER. The results are shown in Fig. 2 after a Monte-Carlo simulation of  $10^6$  symbols for each data point, randomly generated according to Eq. (1) with complex Gaussian  $N$  and wrapped Gaussian  $\varphi$ . It can be seen that for all values of  $\alpha$  and QAM orders (4, 8, 16, and 32), both Foschini's ( $\times$ ) and the proposed approximation ( $\circ$ ) exhibit excellent agreement with the optimal detector (solid line); the Euclidean slicer is also shown for reference (dashed).

In a second simulation, we evaluate the information loss  $\Delta I$  between the exact ( $I_{\text{opt}}$ ) and approximated (or auxiliary,  $I_{\text{aux}}$ ) channel models, in bits per symbol, that is  $\Delta I = I_{\text{opt}} - I_{\text{aux}}$ , where we obtain the channel AIR as [5, Eq. 6]

$$I = m + \frac{1}{M} \sum_{i=1}^M \iint_{\rho, \theta} p(y|x_i) \log_2 \frac{q(y|x_i)}{\sum_{j=1}^M q(y|x_j)} d\rho d\theta \quad (19)$$

where  $M$  is the size of  $\mathcal{X}$ ,  $m = \log_2 M$ , and where we have introduced the distinction between  $p(y|x)$  versus  $q(y|x)$  as in [10, Eq. 35] to enable the computation of both the optimal and the auxiliary-channel lower bounds. Thus, while  $p(y|x)$  is obtained as in Eq. (8) in both cases,  $q(y|x)$  is obtained with Eq. (8) for the optimal case, and with Eqs. (12) and (17) for the approximated channel models under evaluation. Additionally, to ensure that the approximations in Eqs. (12) and (17) comply with the condition  $\int_{\mathcal{Y}} q(y) dy = 1$  we need to apply a normalization of the form

$$\overline{q(y|x_i)} = \frac{q(y|x_i)}{\int_{\mathcal{Y}} q(y|x_i) dy}, \forall i \in \{1, \dots, M\}, \quad (20)$$

Fig. 3 shows the obtained  $\Delta I$  curves for the same values of  $\alpha$  and the same constellations. In the case of AWGN-only ( $\alpha = \infty$ ) both Foschini's metric ( $\times$ ) and the proposed model ( $\circ$ ) show identical behavior, with expected high penalties

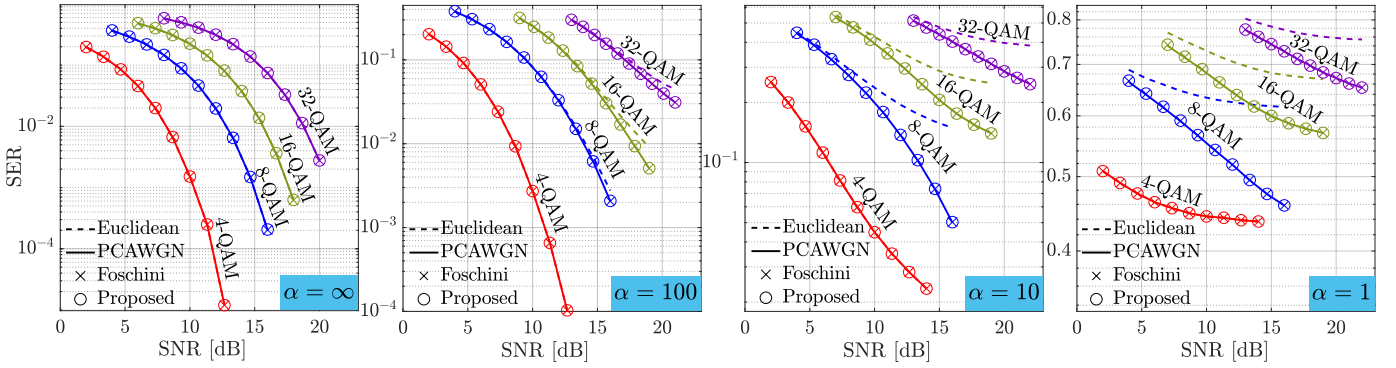


Fig. 2. Symbol error rate (SER) comparison using the models as maximum-likelihood hard-decision detectors (ML-HD). Results shown for the Euclidean detector (dashed), the exact model (solid), Foschini's metric ( $\times$ ), and the proposed approximation ( $\circ$ ).  $\alpha$  takes the values (from left to right):  $\infty$ , 100, 10, and 1. The results are shown for four QAM constellations of order 4, 8, 16, and 32. Each data point corresponds to the simulation of  $10^6$  symbols.

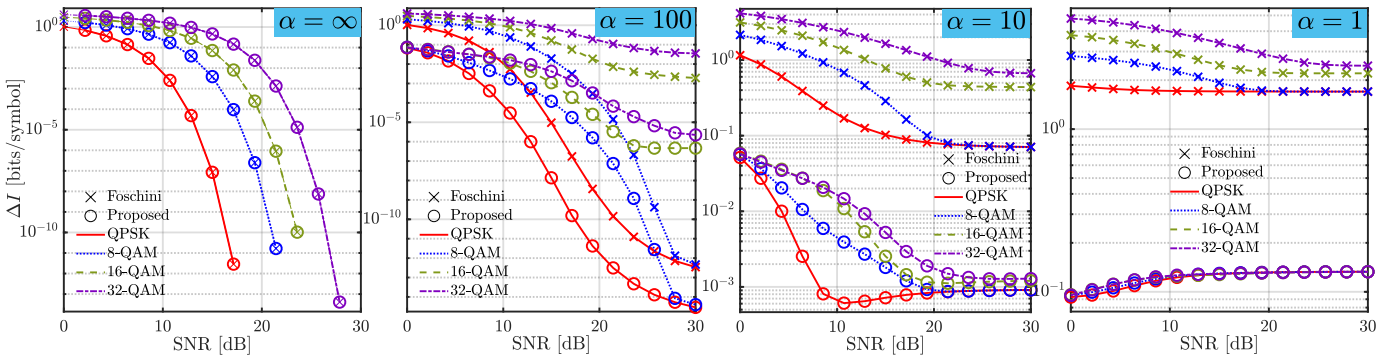


Fig. 3. Information loss ( $\Delta I$ ) in bits/symbol vs. SNR computed as the difference in AIR of the exact channel model (Eq. (8)) and the approximations by Foschini ( $\times$ ) and the proposed model ( $\circ$ ).  $\alpha$  takes the values (from left to right):  $\infty$ , 100, 10, and 1. For each case, the results are shown for four QAM constellations of order 4 (solid line), 8 (dotted line), 16 (dashed line), and 32 (dash-dotted line).

at low SNR and proportional to the constellation order, but quickly vanishing as SNR increases. In contrast, in the presence of phase noise the proposed model exhibits a much better performance when compared to Foschini's metric, reducing the information loss by at least one and up to several orders of magnitude for all the tested constellations and values of  $\alpha$ .

## V. CONCLUSIONS

We have considered the partially coherent AWGN channel (PCAWGN) providing the derivation of the exact solution for the channel conditional probability. We have also suggested ways to approximate this solution, including Foschini's metric, and have introduced a reduced-complexity approximation in polar coordinates that requires less multiplications than the other approaches considered. Our simulations show that both metrics exhibit perfect agreement with the optimum solution as maximum-likelihood hard-decision symbol detectors. However, when computing the channel achievable information rate, our model yields a much lower figure of information loss w.r.t. the exact solution, and reducing by up to three orders of magnitude the values obtained by Foschini's metric.

## REFERENCES

[1] G. J. Foschini, R. D. Gitlin, and S. B. Weinstein. "On the Selection of a Two-Dimensional Signal Constellation in the Presence of Phase Jitter and Gaussian Noise", *The Bell System Tech. Journal*. 52, no. 6 (1973)

[2] R. Corvaja and S. Pupolin. "Phase Noise Effects in QAM Systems", In *Proceedings of 8th International Symposium on Personal, Indoor and Mobile Radio Communications - PIMRC'97*, p. 452-456, vol.2, (1997)

[3] B. Goebel, R. Essiambre, G. Kramer, P. J. Winzer, and N. Hanik. "Calculation of Mutual Information for Partially Coherent Gaussian Channels With Applications to Fiber Optics", *IEEE Transactions on Information Theory* 57, no. 9 (2011)

[4] P. Hou, B. J. Belzer, and T. R. Fischer. "Shaping Gain of the Partially Coherent Additive White Gaussian Noise Channel", *IEEE Communications Letters* 6, no. 5 (2002)

[5] A. Alvarado, T. Fehenberger, B. Chen, and F. M. J. Willems. "Achievable Information Rates for Fiber Optics: Applications and Computations", *Journal of Lightwave Technology* 36, no. 2, p. 424-439 (2018)

[6] M. Sales-Llopis, Md S. Faruk, and S. J. Savory. "Improved Linewidth Tolerant Carrier Phase Recovery Based on Polar MAP Metric Estimate", In *Optical Fiber Communication Conference (2017)*, Paper Th4C.2, Optical Society of America, (2017)

[7] S. O. Rice, "Statistical Properties of a Sine Wave plus Random Noise." *The Bell System Technical Journal* 27, no. 1, p. 109-157 (1948)

[8] H. Fu and P. Y. Kam. "Exact Phase Noise Model and Its Application to Linear Minimum Variance Estimation of Frequency and Phase of a Noisy Sinusoid.", In *2008 IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications*, 1-5, (2008)

[9] K. V. Mardia and P. E. Jupp. *Directional Statistics*. Wiley Series in Probability and Statistics. Chichester; New York: J. Wiley, (2000).

[10] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavcic, and W. Zeng. "Simulation-Based Computation of Information Rates for Channels With Memory". *IEEE Transactions on Information Theory* 52, no. 8 (2006)